

Nonspatial intermodal selective attention is mediated by sensory brain areas: Evidence from event-related potentials

DURK TALSMA AND ALBERT KOK

Psychonomics Department, University of Amsterdam, The Netherlands

Abstract

The present study focuses on the question of whether inter- and intramodal forms of attention are reflected in activation of the same or different brain areas. ERPs were recorded while subjects were presented a random sequence of visual and auditory stimuli. They were instructed to attend to nonspatial attributes of either auditory or visual stimuli and to detect occasional target stimuli within the attended channel. An occipital selection negativity was found for intramodal attention to visual stimuli. Visual intermodal attention was also manifested in a similar negativity. A symmetrical dipole pair in the medial inferior occipital areas could account for the intramodal effects. Dipole pairs for the intermodal attention effect had a slightly more posterior location compared to the dipole pair for the intramodal effect. Auditory intermodal attention was manifested in an early enhanced negativity overlapping with the N1 and P2 components, which was localized using a symmetrical dipole pair in the lateral auditory cortex. The onset of the intramodal attention effect was somewhat later (around 200 ms), and was reflected in a frontal processing negativity. The present results indicate that intra- and intermodal forms of attention were indeed similar for visual stimuli. Auditory data suggest the involvement of multiple brain areas.

Descriptors: Nonspatial attention, Intermodal attention, ERPs, Scalp topography, Source localization

Selective attention can be seen as a collection of hierarchical filters (see Hansen & Hillyard, 1983; Heslenfeld, 1998) that provide various levels of selection. The first of these levels selects information among the sensory modalities, such as the visual, auditory, or tactile modality. At the dimensional level, a given stimulus dimension, such as spatial location, color, orientation (visual modality) pitch or intensity (auditory) is selected from the relevant modality, and at the feature level one particular feature, such as a given color (e.g., the color purple) is selected from the currently selected dimension. Based on this model, one could argue that selection of a relevant modality would precede the selection of relevant stimuli within the attended modality.

Recent studies have questioned such a strict hierarchy, however. It has been shown, for example, that the judgment of the location of auditory stimuli is improved by the presence of a structured visual field (Platt & Warren, 1972), whereas discrepant visual information causes gross errors in auditory location (a phenomenon known as the “ventriloquist” effect; Pick, Warren, & Hay, 1969). In addition, Eimer and Schröger (1998) have demon-

strated the presence of location relevance for both visual and auditory stimuli, which not only existed when presented to the attended modality, but also when these stimuli were presented to a modality which was not attended (i.e., visual stimuli presented when attention was directed to auditory stimuli).

The present article focuses on the question of whether or not nonspatial forms of intermodal attention (as opposed to spatial attention processes discussed above) follow a hierarchical model of selection, and will also attempt to identify which brain areas reflect attentional selection processes. It has been suggested that the attentional systems of the brain are anatomically separate from the data-processing systems (Posner & Petersen, 1990). The anatomical areas identified in selective attention include the posterior parietal lobe (Mountcastle, 1978; Wurtz, Goldberg, & Robinson, 1980), the pulvinar nucleus of the thalamus (LaBerge, 1990; Petersen, Robinson, & Morris, 1987), and the superior colliculus (Posner & Petersen, 1990). It has been shown, however, that attentional systems interact with the operation of data-processing systems. For example, LaBerge (1990, 1995) has argued that thalamic nuclei are able to selectively increase the sensitivity of sensory areas in the visual system. Although LaBerge’s model explicitly discusses the thalamic influence on the visual system, evidence exists for a similar relation between thalamic filtering and the sensitivity of the auditory cortex for processing auditory stimuli (Bastiaansen, Brunia, & Böcker, 1999). A consequence of the neuroanatomical model presented above is that effects of attentional selection are likely to be reflected in a differential activity originating from the visual and auditory sensory areas.

We wish to thank Henric Brouwer and Jasper Hage for assistance in collecting the data, Bert Molenkamp for technical assistance, and two anonymous reviewers for helpful comments on an earlier version of this manuscript.

Address requests for reprints to: Durk Talsma, Psychonomics Department, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands. E-mail: talsma@psy.uva.nl.

Event-related potential (ERP) studies on visual nonspatial attention typically report two different components following attended stimuli. At the occipital sites, the ERPs to attended stimuli usually yield an enhanced negativity, starting around 200 ms after stimulus onset. This component is known as the selection negativity. In addition, at the frontal sites, the ERPs to attended stimuli are shifted positively around 150 to 200 ms (Heslenfeld et al., 1997; Kenemans et al., 1993; Van der Stelt et al., 1998).

Auditory selective attention is manifested in a different pattern of ERP waveforms. Both attended and unattended stimuli elicit prolonged negative potentials, known as "processing negativity" (Näätänen, 1992). The enlarged negativity elicited by attended stimuli has also been termed negative difference (Nd) wave. The Nd consists of an early phase (Nde) with a central scalp distribution that overlaps with the N1, and a late phase (Ndl) that shows a more frontocentral scalp distribution (Hillyard, Mangun, Woldorff, & Luck, 1995).

The late phase of the processing negativity is thought to reflect the matching of a stimulus against an internal template of the relevant stimulus (Näätänen, 1982, 1992), which is known as the "attentional trace." This matching process starts at the same time for both attended and unattended stimuli, but ends earlier for unattended stimuli relative to attended stimuli. Moreover, when the attended and unattended stimuli are relatively easy to discriminate, the processing negativity terminates relatively early compared to stimuli that are relatively hard to discriminate, resulting in an earlier onset in the difference wave (Nd) between relevant and irrelevant stimuli.

Currently, a number of studies have addressed the ERP effects of intermodal selective attention (Alho, Woods, & Algazi, 1994; Alho, Woods, Algazi, & Näätänen, 1992; De Ruiter, Kok & van der Schoot, 1998; Eimer & Schröger, 1998; Hackley, Woldorff, & Hillyard, 1990; Woods, Alho, & Algazi, 1992). A key issue in these studies was the question of whether attentional operations engage supramodal mechanisms or mechanisms that are modality specific. In these studies, subjects typically receive mixed streams of visual and auditory stimuli and are instructed to attend to stimuli presented in one modality, and ignore stimuli presented in the other modality. A common finding was that visual stimuli (standards) presented in the relevant modality yield an enhanced early occipital positivity, peaking around 100 ms (P120/P105), that is followed by a broader negative displacement, which overlaps with the N1 and P2 components (Hackley et al., 1990; Woods et al., 1992, experiment 2). These results were consistent with the existence of neural generators in the same higher order visual areas that also underlie intramodal attention. For auditory standard stimuli, the same studies established that intermodal attention was accompanied by a biphasic modulation (enhancement) of early components (a frontocentral N1 followed by a P2). This pattern contrasts with the typical monophasic processing negativities that are usually observed in intramodal auditory attention.

The suggestion from these early studies is that intermodal attention depends primarily on neural processes in modality-specific cortical areas and not on a single, higher order attentional system. With respect to intermodal auditory attention, ERP results are consistent with the existence of neural generators in the auditory cortex along the superior temporal plane (Woods et al., 1992). For intermodal nonspatial visual attention, the available ERP evidence suggests that attention operates on the basis of the same mechanisms that control intramodal nonspatial attention.

The present study aims at identifying both the time course and scalp distributions of intramodal and intermodal nonspatial attention. To do so, we used an intermodal variant of the filter para-

dig, in which subjects were specifically instructed to direct their attention to a combination of sensory modality and stimulus feature. Similar to Eimer and Schröger (1998), an adapted arrangement for presentation of stimuli was used to ensure that visual and auditory stimuli were presented at approximately the same positions in space. Note that in most previous studies on intermodal attention, auditory stimuli were presented over headphones and visual stimuli were presented from a computer display in front of the subjects, which may have introduced a confound between nonspatial intermodal attention with attention to positions in space (see also Eimer & Schröger, 1998, and Spence & Driver, 1997, for a discussion of this problem). Thus the present design eliminated this confound between attention to modality and spatial attention. Finally, to gain more precise information about the neural origin of these attention effects, the topographical distribution of the ERP waveforms is examined in greater detail by using both scalp topographies and dipole source analyses.

To summarize, the present study addresses the following two questions: (a) Are the brain areas involved in intermodal attention the same as those involved in intramodal attention? (b) Is there a difference in the relative onset of intermodal and intramodal attention? To answer these questions, ERPs were recorded to mixed streams of visual and auditory stimuli. The visual stimuli consisted of two types of horizontal square wave gratings, differing with respect to their spatial frequency. Auditory stimuli consisted of tone bursts with a high and low pitch. Subjects were instructed to selectively attend to features within a certain modality, and detect occasional target stimuli, which were presented with a slightly longer duration, within the attended channel. Thus, the present design allowed us to measure the ERP responses to both visual and auditory stimuli at two levels of relevance, namely, (a) relevance of stimulus features within a modality (intramodal attention) and (b) relevance of the modality of stimuli (intermodal attention).

Based on a hierarchical filtering model of attention, it can be expected that the selection of the relevant modality precedes selecting the relevant feature within the attended modality. However, LaBerge's neurophysiological model of selective attention predicts that a thalamic enhancement circuit selectively enhances the sensitivity of specific groups of neurons in sensory brain areas. That being the case, one might expect that effects of selective attention will originate in specific sensory areas, which contain large groups of neurons that are responsive to the attended stimulus feature. If the latter is the case, it can be expected that both nonspatial intramodal and intermodal attention effects will be found at similar latencies, but that amplitudes of affected ERP components will be modulated as a function of various levels of relevance of a given stimulus.

Method

Participants

Twenty healthy people participated in the experiment (age 18–41, mean 22; 9 men and 11 women). All participants had normal or corrected-to-normal vision and normal hearing capabilities. All participants took part in the experiment as part of a requirement for their introductory course in psychology.

Task and Stimuli

Streams of two different visual and auditory stimuli were presented in randomized order to the participants. The visual stimuli consisted of white rectangular horizontal square wave gratings, with a spatial frequency of 3.2 cycles per degree for the high spatial frequency gratings and 0.8 cycles per degree for the low spatial fre-

quency. The values used are known to elicit the largest differences in their exogenous effects (Previc & Harter, 1982). The auditory stimuli consisted of sine waves, with a frequency of 900 Hz for the low pitched tones and 2000 Hz for the high pitched tones. Tones were presented at 65 dBA, and had linear rise and fall times of 10 ms. Regular nontarget gratings and tones (see below for a distinction between targets and nontargets) were presented randomly with equal probability and lasted for 50 ms. Visual stimuli were presented at the central location (point of fixation) at a computer screen, and auditory stimuli were presented over two loudspeakers that were mounted to the lateral edges of the screen. This procedure minimized the use of spatial information about the location of relevant and irrelevant stimuli (see also the apparatus section below).

Stimuli were presented in a randomized order, in attend-visual and attend-auditory blocks lasting about 6 min and each containing 280 visual and 280 auditory stimuli. Prior to each experimental block, participants were instructed to attend to stimuli presented to either the visual or the auditory modality. In addition, they were required to attend only to one stimulus type in the attended modality that contained the relevant stimulus attribute (spatial frequency for visual and pitch for auditory stimuli) and respond by pressing a button whenever a target stimulus (25% of the trials) was presented within the attended channel. Target stimuli were equal to standard stimuli, but were presented with a slightly longer duration (visual stimuli: 200 ms, auditory stimuli: 100 ms). A longer duration of visual stimuli was selected because pilot studies had shown that the visual target/nontarget differences were very difficult to discriminate at shorter durations. It should be noted that the longer duration stimuli were not only present in the relevant stimulus category, but in the irrelevant as well. That is, when subjects were instructed to attend to low frequency gratings, 25% of those gratings were actually targets. In addition, 25% of the other stimuli were also presented with a longer duration.

The task was presented in two conditions, hereafter referred to as the single modality condition and the mixed modality condition. In the single modality condition, participants received only visual or auditory stimuli (280 per run; total duration about 4.5 min), whereas in the mixed modality condition, all the stimuli were presented (560 per run; total duration about 6 min). The interstimulus interval (ISI) varied randomly between 417 and 817 ms in the mixed condition and between 417 and 1,234 ms in the single modality condition.¹

Throughout the remainder of this article, the following operational definitions will be used for the three levels of attention: "attended frequency" stimuli indicate stimuli of both attended feature (spatial frequency or pitch) and modality (visual or auditory). "Unattended frequency" stimuli indicate those stimuli that consisted of the irrelevant feature, but were presented to the relevant (attended) sensory modality. Finally, the term "unattended modality" stimuli is used to indicate those stimuli that were presented to the unattended modality.

Apparatus

The visual stimuli were presented centrally on a 21-in. VGA computer display, located at a distance of 56 cm, directly in front of the subjects' eyes. Auditory stimuli were presented simultaneously over two loudspeakers that were placed at the two sides of

the display, at an eccentricity of about 15 deg from the center of the screen. This arrangement created the impression that the tones came from a position approximately in the middle of the display, at the point of fixation. Stimulus presentation was controlled by a 100-MHz pentium personal computer, running an in-house developed MS-DOS application to ensure exact timing.

The EEG was continuously recorded using 30 electrodes mounted in an electrocap. The locations used were Fpz, AFz, Fz, Cz, Pz, Oz, Fp1, Fp2, F7, F3, F4, F8, FC5, FC6, T7, C3, C4, T8, CP5, CP1, CP2, CP6, P7, P3, P4, P8, PO3, PO4, O1, and O2. The electrodes were referenced against an electrode attached to the participants' right ear lobe. The impedance was kept below 5 k Ω . Horizontal eye movements were measured by deriving the electro-oculogram (EOG) from two electrodes placed to the outer canthi of the subjects' eyes. Vertical eye movements and eye blinks were detected by deriving an EOG from two electrodes placed approximately 1 cm above and below the subject's right eye. The EEGs and EOGs were fed to a Nihon-Kohden Neurotop amplifier and bandpass filtered using a 35-Hz low-pass filter and a time constant of 2.5 s. The registered signals were digitized with a sample frequency of 250 Hz and stored on CD-ROM discs for off-line analysis.

Procedure

The single and mixed modality conditions were presented in a randomized order across participants and the four attention conditions were also presented randomly in separate blocks. These conditions were: attend-auditory high-pitch tones, attend-auditory low-pitch tones, attend-visual high-frequency gratings, and attend-visual low-frequency gratings. Each condition was presented twice to each participant. A new stimulus order was randomly generated before each block.

After attachment of the electrodes, participants were given a number of practice trials, to ascertain they understood the paradigm and to familiarize them with the stimulus material. Before each run, the participants were instructed to attend to stimuli that contained the relevant modality and feature. (e.g., high-pitched tones) and to respond only to the target stimulus (e.g., the longer duration high-pitched tones) and to respond as fast as possible to those targets while avoiding making too many errors. They were further instructed to minimize the number of eye blinks, to keep their eyes fixated on a centrally presented fixation cross, and to move as little as possible. The eye movements were monitored using both the EOG and a TV camera to verify that the subjects maintained fixation in the attend-auditory conditions.

Data Analysis

During off-line analysis, time-locked epochs of 1,024 ms (256 samples) containing a prestimulus baseline of 100 ms that contained no amplifier saturations, horizontal eye movements, or artifacts (50- μ V amplitude deviation per 4 ms) were selected and stimulus-locked averaged. Any remaining ocular artifacts were corrected (e.g., Kenemans, Molenaar, & Verbaten, 1991). All averages were corrected for possible overlap between successive trials using the ADJAR level 2 method (Woldorff, 1993).²

¹These ISIs were chosen to ensure that the distribution of time intervals between successive stimuli in the single modality condition were almost identical to the distribution of intervals that separated consecutive visual (or auditory) stimuli in the mixed modality condition.

²Although the amount of trial overlap is not very large in the present study, during the shortest ISIs, trials partially overlap the negative-going slope of the P300, resulting in a small but consistent negative slope during baseline and the initial phase of mainly the posterior visual ERPs. A comparison of the original and ADJAR corrected ERPs showed small but consistent changes in the slope of the initial phase of the ERP, with the above-mentioned slope being greatly reduced.

ERP responses to high and low frequency gratings (attend-visual conditions) and high and low tones (attend-auditory conditions) were collapsed. Thus, condition-wise averaging of stimuli presented to the attended modality in both the single and mixed modality conditions was based on a combination of two categories. Moreover, in the mixed modality condition, the responses to visual stimuli in the attend auditory sessions consisted of a combination of high and low spatial frequency stimuli obtained during the attend auditory-low frequency and attend auditory-high frequency sessions. In this case, averaging was based on a total of four categories. The same approach was used for averaging the auditory responses during the attended-visual sessions in the mixed modality condition. Notice that averaging was based only on standard (nontarget) stimuli, and that targets were excluded from the analysis.

Dipole Analysis

Dipoles were fitted using the BESA software package (Scherg & Berg, 1995), using difference waves resembling the attention effects as input for this program. To obtain reliable source estimations, the following procedure was used. First, the localization procedure was limited to latency ranges for which statistically significant attention effects were found. Secondly, for these latencies, the estimated residual variance was plotted as a function of time. If these plots showed sharp decreases of residual variance in a given subwindow, this information was used to further restrict the time windows for which dipoles were fitted. Thirdly, during the localization phase, dipole fits were obtained by using various starting points, to ensure that the localization procedure was not prematurely ended in a local pool of minimum residual variance. Fourth, for reasons of both biological plausibility and mathematical simplicity, only symmetrical dipole pairs were used. The latter is also important because dipole solutions that require the estimation of a large number of free parameters tend to become unstable when they are based on the relatively low number of electrodes used in the present study. As shown by Huizenga and Molenaar (1994), the localization error due to estimating six or seven free parameters (that is one single dipole or two symmetrical dipole pairs) is still relatively low when based on 30 to 40 channels. Thus, for the relatively simple dipole models we used, the localization error is much less a problem and previous studies have been successful in estimating relatively simple dipole models using about 30 channels (e.g., Heslenfeld et al., 1997; Kenemans, Baas, Mangun, Lijffijt, & Verbaten, 2000; Lange, Wijers, Mulder, & Mulder, 1998, 1999; Wijers, Lange, Mulder, & Mulder, 1997; Wijers, van Hooff, Lange, Peters, & Dunajski, 1993).

Results

Behavioral Measurements

Average response times for the correct responses to target stimuli and the percentage of misses and false alarms were calculated for each condition and subjected to an SPSS-MANOVA analysis with the following within-subjects factors: Modality (two levels: visual vs. auditory), Frequency (two levels: high vs. low frequency), and Task (two levels: single modality vs. mixed modality).

Table 1 provides an overview of the observed average response times and percentages of misses and false alarms, with standard deviations for these measures. Statistical analysis revealed significant effects for Modality on response times, $F(1, 19) = 27.76$, $p < .000$, percentage of false alarms, $F(1, 19) = 7.33$, $p < .014$, and percentage of misses, $F(1, 19) = 19.93$, $p < .000$. These results

Table 1. Mean Response Times, Percentage Misses and False Alarms

	Visual stimuli		Auditory stimuli	
	Attend left	Attend right	Attend left	Attend right
Response times				
Mixed	501 (67.4)	515 (66.4)	505 (83.7)	515 (74.0)
Single	504 (66.3)	516 (71.7)	491 (45.7)	515 (60.8)
False alarms				
Mixed	1.90 (1.97)	2.50 (2.74)	3.35 (6.75)	2.52 (2.19)
Single	1.63 (1.94)	1.65 (1.62)	1.92 (2.07)	1.73 (2.58)
Misses				
Mixed	10.6 (11.0)	12.1 (10.7)	12.9 (21.5)	9.28 (5.70)
Single	10.9 (9.28)	13.2 (13.3)	14.9 (14.5)	8.65 (5.62)

Note: Standard deviations are given in parentheses.

indicate that the auditory conditions were somewhat more difficult than the visual conditions. No statistical evidence was found, however, that indicated behavioral differences between the single and mixed modality task conditions, which can be taken as evidence that the single and mixed modality conditions were about equally difficult.

Visual ERPs

Visual ERPs consisted of an early occipital N1, which was followed by P2, N2, and posterior P3 component. Statistical analysis was performed using an SPSS-MANOVA within-subjects design. The Oz electrode was chosen for this analysis after visual inspection of the averaged ERP waveforms and their scalp topographies, which revealed that the largest effects were found over this electrode. These analyses contained the factor Attention (two levels: attended frequency and unattended frequency; attended frequency and unattended modality; and unattended frequency and unattended modality, depending on the specific comparison being analyzed). Separate MANOVA analyses were run on successive averaged time windows (of 12 ms each) in the latency range between 0 to 400 ms. Because of the high number of tests (and to minimize statistical bias), effects are reported only when an effect became significant ($\alpha < .01$ two-tailed) for at least three consecutive samples.

Single Modality Condition. As can be seen in Figure 1, attended frequency stimuli elicited larger negativities, which became statistically significant between 204 and 324 ms after stimulus onset. See Table 2 for a detailed overview of statistical results.

Mixed Modality Condition. A similar intramodal pattern of results was found in the mixed modality condition, which is shown in Figure 2. An SPSS-MANOVA analysis revealed that the negativity caused by intramodal attention was significant from 216 to 360 ms after stimulus onset (see Table 2).

Both attended frequency and unattended frequency stimuli elicited enhanced negativities over the occipital areas, relative to unattended modality stimuli, in the latency range of about 180 to 300 ms after stimulus onset. As can be seen in Table 2, ERPs elicited by both attended frequency and unattended frequency stimuli deviated from unattended modality stimuli. Note that this effect became significant earlier (and remained significant longer)

Visual ERPs, Single Modality

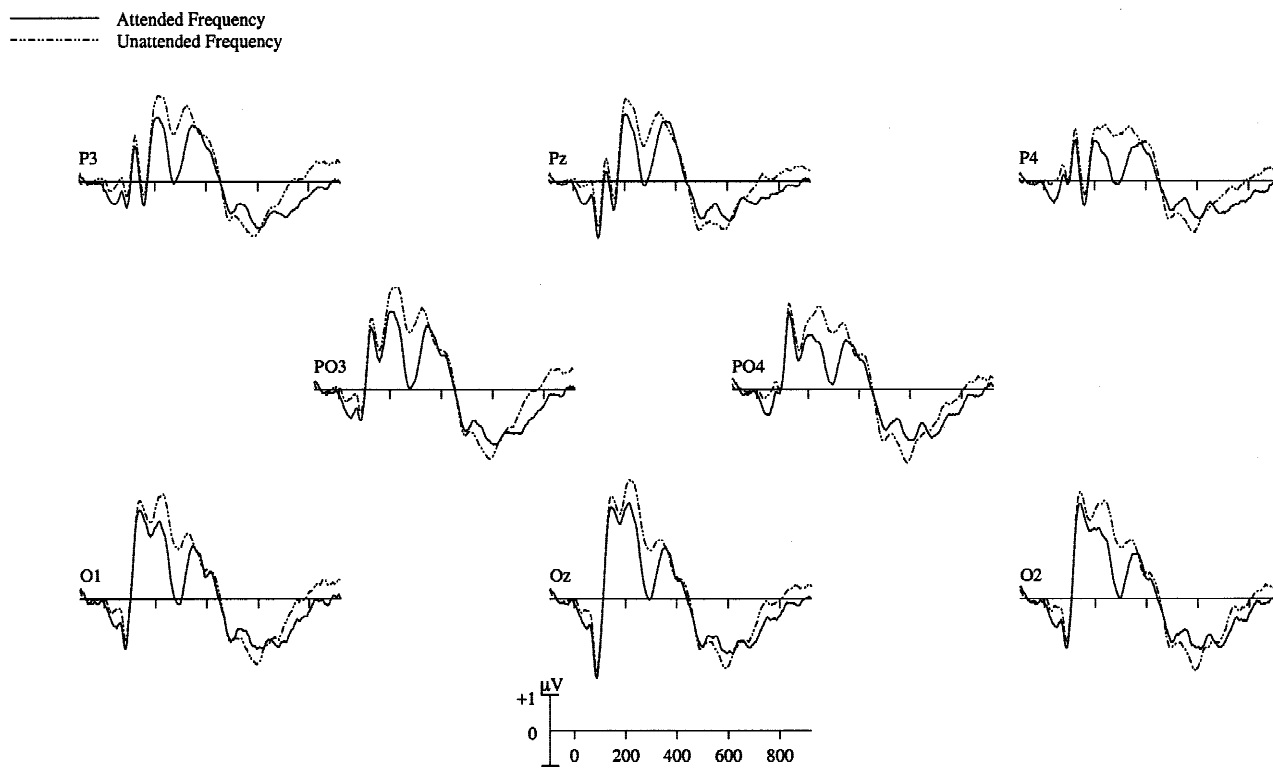


Figure 1. Grand average visual ERP waveforms from a selection of eight posterior channels, obtained in the single modality condition. Attention related effects are clearly visible from about 200 to 300 ms after stimulus onset, as an enhanced negativity for attended frequency stimuli.

for the attended frequency versus unattended modality contrast than unattended frequency versus unattended modality contrast. In addition, Table 2 also reveals significant effects of intermodal attention on later latency ranges (after about 324 ms), in which interval the unattended modality ERPs show a more negative-going slow wave, as compared to attended frequency and unattended modality ERPs.

Differences Between Single and Mixed Modality. As can be seen by comparing Figures 1 and 2, between 108 and 264 ms after stimulus onset, the P2 components are somewhat larger in the single modality condition than in the mixed modality condition. This observation is statistically evidenced by significant effects of the factor Task in this time window. This main effect of Task does not, however, influence the observed differences on the attention effects, as is evidenced by the absence of significant Task \times Attention interactions during this time range.

A main effect of Attention, for the combined single and mixed modality conditions, became significant between 196 and 348 ms after stimulus onset (see Table 2 for statistical details). This latency range is similar to that reported for effects of intramodal visual attention in both the single and mixed modality separately.

No statistically significant Task \times Attention interactions were found throughout the entire time range subjected to the analysis, indicating that the presence of auditory stimuli during a visual attention task did not interfere with visual selection, as reflected in the time course of the ERP wave forms.

Scalp Topographies and Source Localizations. For each significant time window described above, normalized mean amplitudes (McCarthy & Wood, 1985) were calculated per contrast, for each of the 30 EEG channels. These amplitudes were submitted to SPSS-MANOVA analyses with the following within-subjects factors: Channel (30 levels) and Attention (two levels). Greenhouse–Geisser correction was applied to all tests involving the factor Channel. Note that in the present analyses, the F ratios associated with the Channel \times Attention interaction are of specific importance because they index the extent to which relevant attention conditions differ in their overall scalp topographies in the selected time windows. Thus, these results were useful to substantiate visual inspection of the topographical maps provided by the BESA program (see further below).

A source localization was performed on the effects of intra- and intermodal attention, using the BESA software package (Scherg & Berg, 1995). These analyses were restricted to the time windows in which the preceding time-slice analyses had revealed significant effects of intra- and intermodal attention. These analyses were performed on the grand averages. The neural sources were estimated using a symmetrical pair of spatiotemporal dipoles. Furthermore, results of these localizations are reported only for those epochs where a satisfactory residual fit (less than 5%) could be obtained.

Figure 3 shows spline and current source density (CSD) maps and dipoles for the difference wave obtained by subtracting ERPs following unattended frequency stimuli from ERPs following

Table 2. Overview of the Time Slice Analyses for the Effects of Visual Stimuli, as Performed on the Oz Electrode

Time (ms)	Mixed			Single	Single vs. mixed		
	af vs. uf	af vs. um	uf vs. um	ATT	TASK	ATT	TASK \times ATT
72	NS	NS	NS	NS	NS	NS	NS
84	NS	NS	NS	NS	NS	NS	NS
96	NS	NS	NS	NS	NS	NS	NS
108	NS	4.90	9.98	NS	4.93	NS	NS
120	NS	8.07	20.30	NS	13.88	NS	NS
132	NS	6.00	15.54	NS	19.74	NS	NS
144	NS	NS	NS	NS	17.07	NS	NS
156	NS	NS	NS	NS	15.17	NS	NS
168	NS	10.65	NS	NS	16.45	NS	NS
180	NS	19.07	13.79	NS	14.35	NS	NS
192	NS	28.37	29.05	4.67	10.21	6.26	NS
204	NS	38.01	41.89	11.38	6.71	13.80	NS
216	5.05	43.56	45.85	21.51	5.30	26.71	NS
228	9.30	47.75	50.10	25.38	5.56	36.24	NS
240	15.96	49.38	52.56	25.99	5.45	41.69	NS
252	19.99	49.18	51.73	33.45	4.80	44.70	NS
264	22.91	46.84	38.83	47.58	NS	51.33	NS
276	28.61	38.79	21.31	51.71	NS	59.82	NS
288	31.39	28.72	8.93	36.53	NS	50.13	NS
300	27.61	21.95	NS	25.99	NS	34.69	NS
312	20.22	14.67	NS	16.47	NS	21.31	NS
324	14.75	NS	NS	7.24	NS	12.34	NS
336	11.50	NS	10.36	NS	NS	5.73	4.69
348	6.19	NS	22.93	NS	NS	NS	5.55
360	NS	8.43	31.61	NS	NS	NS	NS
372	NS	12.30	25.38	NS	NS	NS	NS
384	NS	10.58	17.66	NS	NS	NS	NS
396	NS	6.74	12.24	NS	NS	NS	NS

Note: All reported F values (1,19 degrees of freedom) are significant with an $\alpha < .01$ (two-tailed). NS: not significant; ATT: factor Attention; TASK: factor Task; af: attended frequency; uf: unattended frequency; um: unattended modality; Mixed: mixed modality condition; Single: single modality condition.

attended frequency stimuli. These were fitted in the 231 to 324 ms interval. Spline and CSD maps show that the intramodal attention effect is associated with a broad negativity with a focus over the secondary (extrastriate) posterior cortical areas. These scalp topographies differed between attended frequency and unattended frequency stimuli, as indicated by a significant Attention \times Channel interaction, $F(29,511) = 4.15$, $p < .008$, $GG = .11$. A symmetrical dipole pair in the medial inferior posterior areas yielded a stable solution, and a residual variance of 2.45%.

Figures 4 and 5 show the topographical maps and dipoles for the difference waves obtained by subtracting ERPs elicited by attended frequency stimuli (Figure 4) and ERPs elicited by unattended frequency stimuli (Figure 5) from ERPs elicited by unattended modality stimuli. The attended frequency minus unattended modality difference wave was fitted in the latency range of 169 to 324 ms and the unattended frequency minus unattended modality difference wave was fitted between 190 and 272 ms after stimulus onset. Both difference waves yielded symmetrical dipole pairs at similar locations in the medial inferior occipital areas, with residual variances of 3.45% and 2.95%, respectively. These areas globally coincide with ventromedial extrastriate visual cortex. Topographical analyses of the scalp distributions during the latency ranges reported above indicate that attended frequency and unattended modality gratings generated different scalp topographies (Attention \times Channel interaction, $F(29,511) = 7.07$, $p < .001$, $GG = .09$). However, evidence for topographical differences between unattended frequency and unattended modality gratings failed to reach significance, $F(29,511) = 2.45$, $p > .05$, $GG = .08$.

Auditory ERPs

Auditory ERPs consisted of frontocentral P1, N1, and P2 components, followed by a small P3, which was followed by a slow negative shift for attended frequency stimuli. For auditory stimuli, a statistical analysis similar to that of the visual stimuli was used. However, inspection suggested that effects of attention were most pronounced at the frontocentral electrode locations, and thus Cz was used as the principal site for statistical analysis.

Single modality condition. Figure 6 shows the auditory ERPs at a selection of frontocentral leads. Most prominently, a processing negativity, which follows the N1 component can be observed.

Mixed modality condition. The mixed modality condition showed a similar pattern of results for intramodal attention. In addition, attended frequency and unattended frequency stimuli also elicited a late negativity (relative to the other stimuli) in the latency range between 300 and 400 ms after stimulus onset.

For intermodal attention, the time course of the results is somewhat more complex (see Figure 7). Intermodal attention started relatively early with a negative modulation of the N1 amplitude, which was followed by a modulation of the P2. The latter effect, however, depended on the type of contrast. The P2 elicited by attended frequency stimuli was smaller, whereas the P2 elicited by unattended frequency stimuli was larger than the P2 following unattended modality stimuli.

The statistical analyses confirmed that the time course of the intermodal effect differed markedly from that of the intramodal

Visual ERPs, Mixed Modality

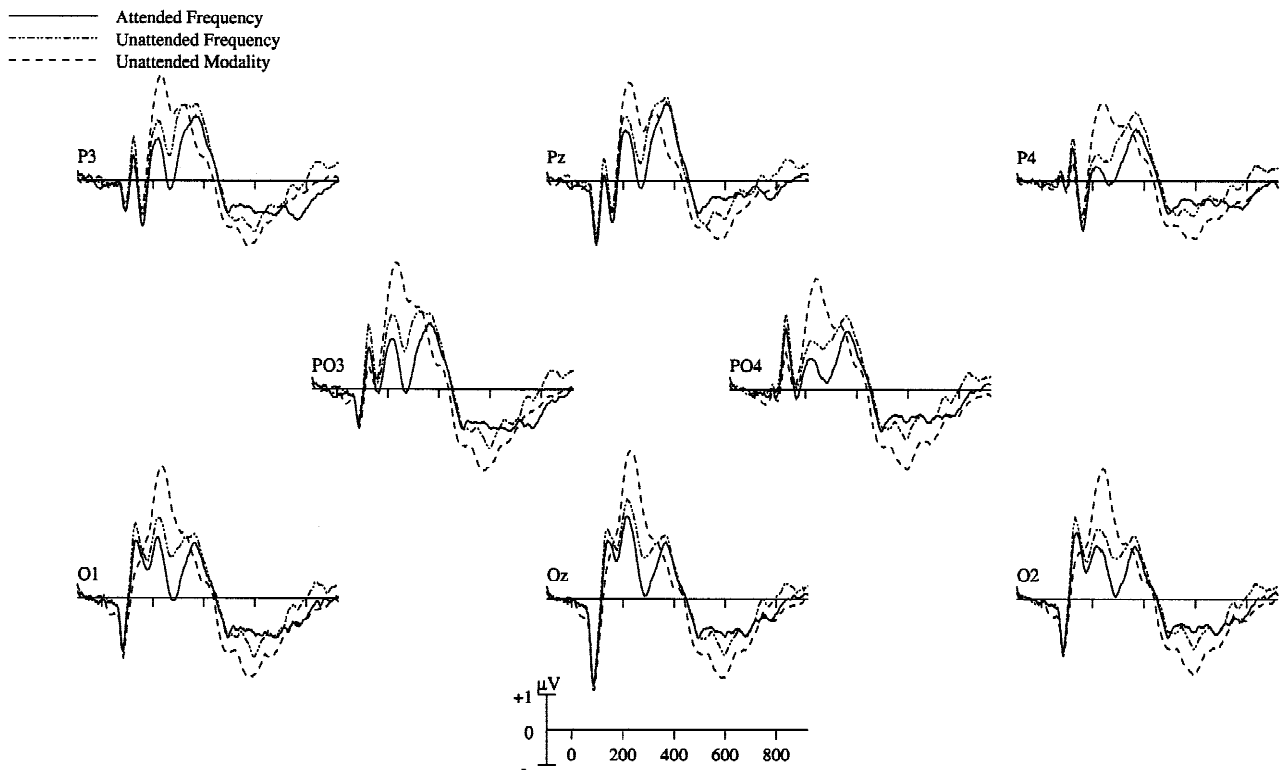


Figure 2. Grand average visual ERP waveforms from a selection of eight posterior channels, obtained in the mixed modality condition. Intermodal attention is visible as an enhanced selection negativity, following both attended frequency and unattended frequency stimuli. As in the single modality condition, intramodal attention is reflected by an increased negativity following attended frequency stimuli, from about 220 to 300 ms after stimulus onset.

effect. As can be seen from the results of the three planned contrasts, (see Table 3), intramodal attention only revealed a significant effect during the second interval previously mentioned (between 204 and 276 ms, globally the P2 region). The two intermodal contrasts, on the other hand, show a somewhat more complex pattern. Whereas the attended frequency versus unattended modality contrasts yielded only significant differences during the early N1 interval, the unattended frequency versus unattended modality contrasts showed significant differences in two intervals.

Differences between single and mixed modality conditions. On one short interval (between 60 to 84 ms after stimulus onset), a main effect of Task was found, indicating that the P1 amplitude is somewhat larger in the single modality condition than in the mixed modality condition. Although a comparison of Figures 6 and 7 also suggests an interaction between Task and Attention for the amplitude of the P1 component, no statistical evidence for such an effect could be found (see Table 3). A main effect of Attention was found between 180 and 288 ms after stimulus onset, which corresponds to the processing negativities that occur in this time region, as can be seen in Figures 7 and 6. One should also notice that the amplitude difference between attended frequency and unattended frequency tones is somewhat larger in the single modality condition than in the mixed modality condition. This observation is statistically confirmed by a significant Task \times Attention interaction between 204 and 228 ms after stimulus onset.

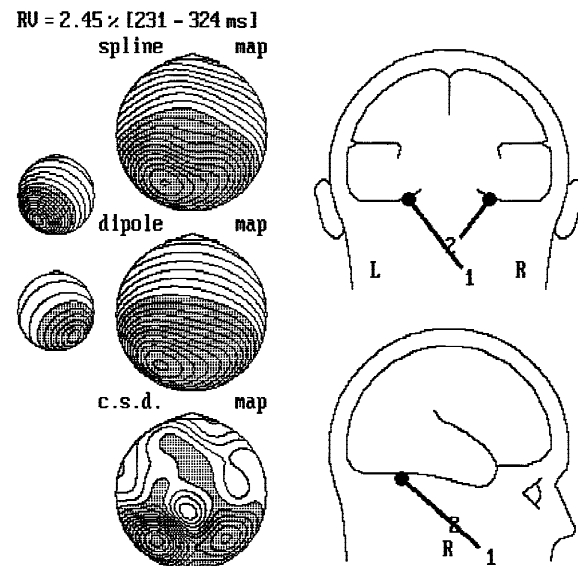


Figure 3. Dipole localization of visual intramodal attention for the mixed modality condition. Contour maps were calculated at 281 ms after stimulus onset. Far left: scalp distribution resulting from dipoles 1 (top) and 2 (bottom). Left: measured scalp topography (top), modeled scalp topography (center), and current source density (CSD) map of the fitted dipoles. Contour spacing is $0.1 \mu\text{V}$ for the spline maps and $0.05 \mu\text{V}/\text{cm}^2$ for CSD maps. Right: coronal (top) and sagittal (bottom) view of the fitted dipole pair.

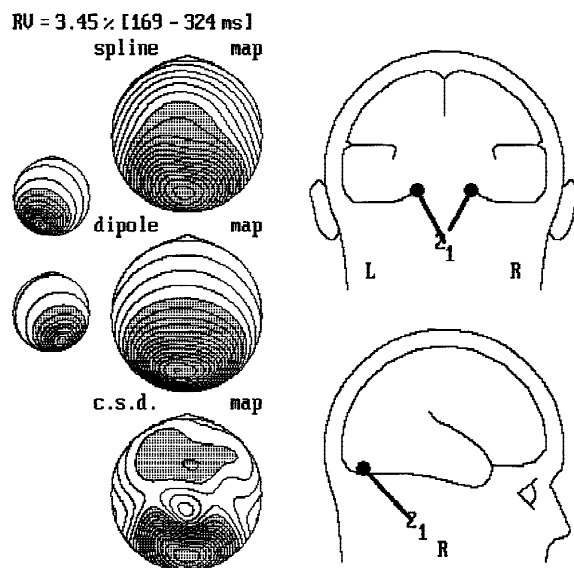


Figure 4. Dipole localization of visual intermodal attention, as fitted on the difference wave obtained by subtracting ERPs following unattended modality stimuli from ERPs following attended frequency stimuli. Contour maps were calculated at 240 ms after stimulus onset. Far left: scalp distribution resulting from dipoles 1 (top) and 2 (bottom). Left: measured scalp topography (top), modeled scalp topography (center), and current source density (CSD) map of the fitted dipoles. Contour spacing is $0.5 \mu\text{V}$ for the spline maps and $0.25 \mu\text{V}/\text{cm}^2$ for CSD maps. Right: coronal (top) and sagittal (bottom) view of the fitted dipole pair.

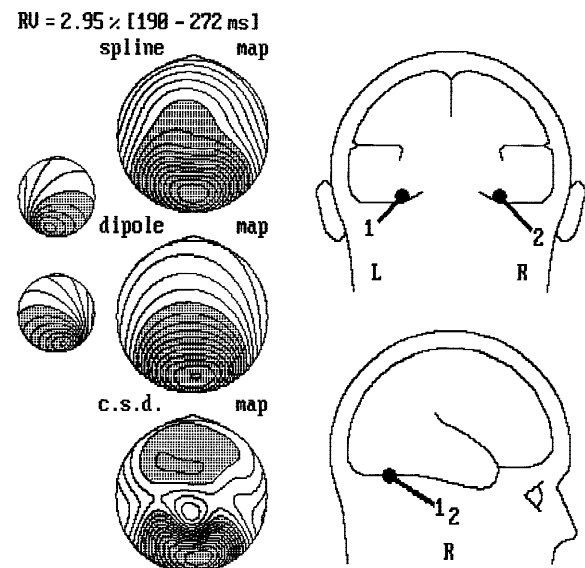


Figure 5. Dipole localization of visual intermodal attention, as fitted on the difference wave obtained by subtracting ERPs following unattended modality gratings from ERPs following attended frequency gratings. Contour maps were calculated at 230 ms after stimulus onset. Far left: scalp distribution resulting from dipoles 1 (top) and 2 (bottom). Left: measured scalp topography (top), modeled scalp topography (center), and current source density (CSD) map of the fitted dipoles. Contour spacing is $0.5 \mu\text{V}$ for the spline maps and $0.25 \mu\text{V}/\text{cm}^2$ for CSD maps. Right: coronal (top) and sagittal (bottom) view of the fitted dipole pair.

Scalp topographies and source localizations. For the auditory stimuli, reasonable dipole fits were obtained only for the two difference waves that expressed intermodal attention effects (i.e., the attended frequency minus unattended modality and unattended frequency minus unattended modality difference waves) in the time window between 96 and 128 ms after stimulus onset. These dipoles correspond to the early amplitude modulation of the N1 component, which was enhanced in amplitude for the attended frequency and unattended frequency auditory stimuli, relative to the unattended modality stimuli.³ The topographical analyses were confined to the scalp distributions generated during the interval for which the neural sources were fitted. The Attention \times Channel interactions obtained from these analyses indicate that (a) ERPs to attended frequency and unattended modality tones did not differ significantly in their overall scalp distributions, $F(29,511) = 2.08$, $p > .10$; $GG = .09$, and (b) that these distributions were somewhat different between ERPs following partially attended frequency and unattended modality stimuli, $F(29,511) = 5.99$, $p < .005$, $GG = .07$.

Figures 8 and 9 show the topographical and dipole analysis results for these two difference waves. Spline and CSD maps of the

waveforms show a clear frontocentral negative distribution of difference waves that were associated with auditory intermodal attention effects.

As can be seen from Figure 8 (attended frequency minus unattended modality difference wave), a symmetrical dipole pair in the lateral anterior temporal areas, fitted in the latency range of 96 to 128 ms, yields a stable solution, with a residual variance of 3.59%. The attended frequency minus unattended modality difference wave (Figure 9) yields a similar result, in the same latency range, with a residual variance of 3.4%.

Figures 10 (attended frequency minus unattended frequency) and 11 (unattended frequency minus unattended modality) show dipole fits and scalp topographies fitted at 231 ms after stimulus onset, during the second time window showing significant effects of auditory attention. Note that during this window, ERPs to unattended frequency stimuli are significantly more positive than ERPs to attended frequency and unattended modality stimuli. As a consequence the topographies shown in Figures 10 and 11 are shown in reverse polarity. Although no statistically significant effects between attended frequency and unattended modality stimuli were found on the time-slice analyses, both the dipole solutions and the scalp topographies are markedly different for the two solutions shown in Figures 10 and 11, which could be taken as an indication that different neural systems are involved in the later stages of intra- and intermodal auditory attention. These observations were confirmed by statistical analysis, which showed that attended frequency and unattended frequency tones did not produce significantly different scalp topographies, $F(29,511) = 3.55$, $p > .01$, $GG = .10$, but that unattended frequency and unattended modality tones did, $F(29,511) = 8.23$, $p < .0001$, $GG = .09$.

³Note that for all auditory dipole fits, the latency ranges over which dipoles were fitted are notably shorter than the time windows that are reported as showing statistically significant effects. The reason for this discrepancy is the following. When the whole latency range showing significant effects was subjected to the dipole analysis, fits were generally poor. Residual variance plots of these fits indicated that smaller subsections of these windows did yield reasonable fits and when the analysis was confined to these sections, adequate fits could be obtained. These subwindows also appear to correspond to those latency ranges where the largest deflection in the corresponding difference wave occurred and where the highest F-ratios were found.

Auditory ERPs, Single Modality

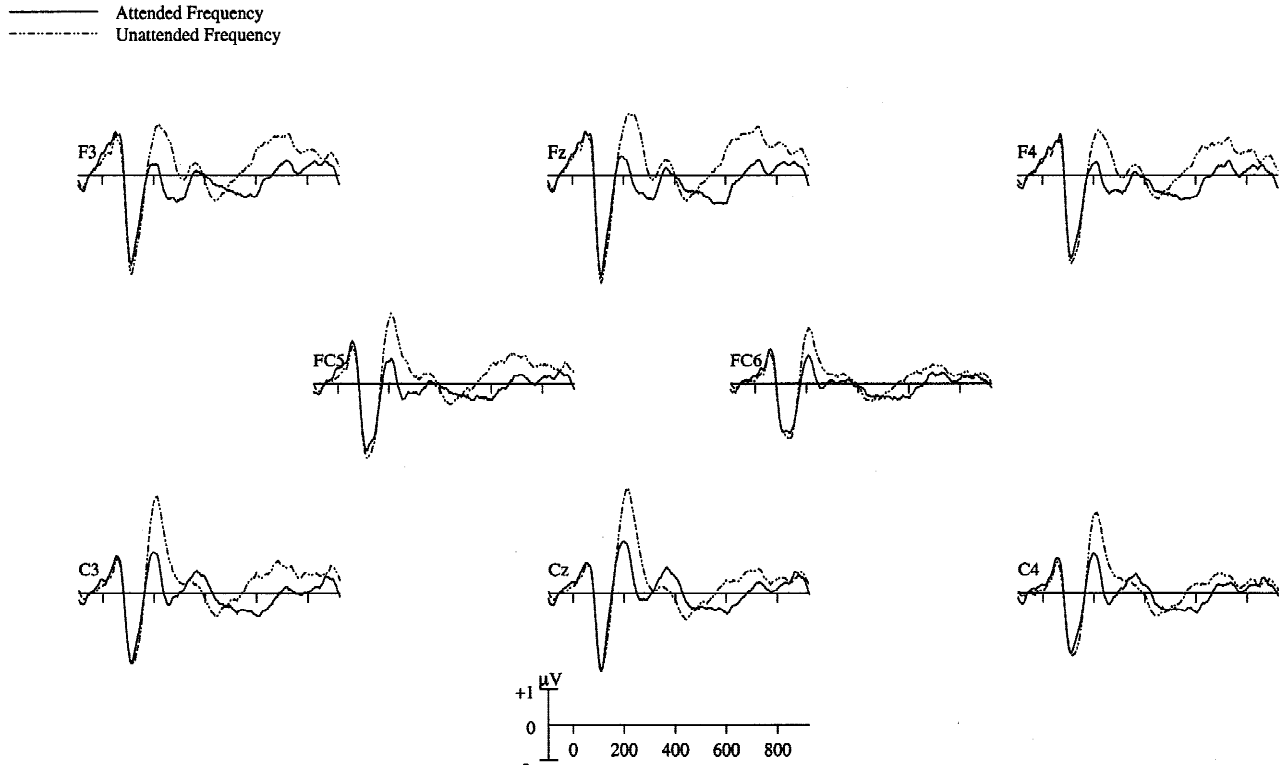


Figure 6. Grand average auditory ERP waveforms from a selection of eight anterior channels, obtained in the single modality condition. An attention related processing negativity starts around 140 ms after stimulus onset.

Summary of Results

The main findings of the present study are that both intermodal and intramodal forms of attention are modulated by sensory brain areas in the visual and auditory cortices. It was found that the onset of the visual intermodal attention effect was somewhat earlier than that of the visual intramodal effect. For both effects, dipoles could be fitted in the occipital areas. The ERP waveforms showed that this effect was manifested in an occipital selection negativity, which overlapped the occipital P2 and N2 components.

For auditory stimuli, a similar pattern of results was found. The onset of the intermodal effect was somewhat earlier than the onset of the intramodal effect. For early intermodal attention, a symmetrical dipole pair in the auditory areas could explain the observed auditory N1 amplitude modulation. Intramodal attention, however, showed a somewhat different pattern of results; here attention effects are mainly reflected in the later part of the auditory processing negativity.

Discussion

In the present study, both intermodal and intramodal manipulations of attention were varied independently, which allowed a direct comparison of these forms of attention in the same experiment. Our study aimed at answering two basic questions with regard to the characteristics of inter- and intramodal attention. The first question is whether attentional operations in intermodal and intramodal conditions are mediated by the same modality-specific attentional systems. The second question addresses the relative time

course of intermodal and intramodal attention. It was hypothesized that intermodal attention effects on the ERP waveforms should have an earlier onset, because stimuli that differ in their modality are physically more distinct than stimuli presented in the same modality.

Visual ERPs

Morphology of the ERP waveforms. The morphology of the raw ERP waveforms was somewhat different, compared to what is commonly reported in the literature. It should be noticed that over the occipital electrodes, no P1 components were observed and that the latencies of the N1 and P2 components are somewhat earlier than what is typically reported. These differences are possibly caused by the relatively high presentation rate. Notice that Helsenfeld (1998, chapter 5) also reported similar waveforms for the same type of gratings as used in the present study, which were also presented at a comparably high presentation rate.

Intermodal effects. Intermodal visual attention was manifested in a negative modulation of ERPs to attended frequency stimuli relative to ERPs elicited by visual stimuli when audition was the relevant modality. These effects strongly resembled those that were obtained in the earlier study by De Ruiter et al. (1998) and effects reported by Alho et al. (1992). Woods et al. (1992) found similar results as well, while using lateralized stimuli. In the latter studies, ERPs to visual standards consisted of an N1 and P2 deflection that showed an enhanced negativity at the occipital electrodes relative to ERPs elicited in the attend auditory condition.

Auditory ERPs, Mixed Modality

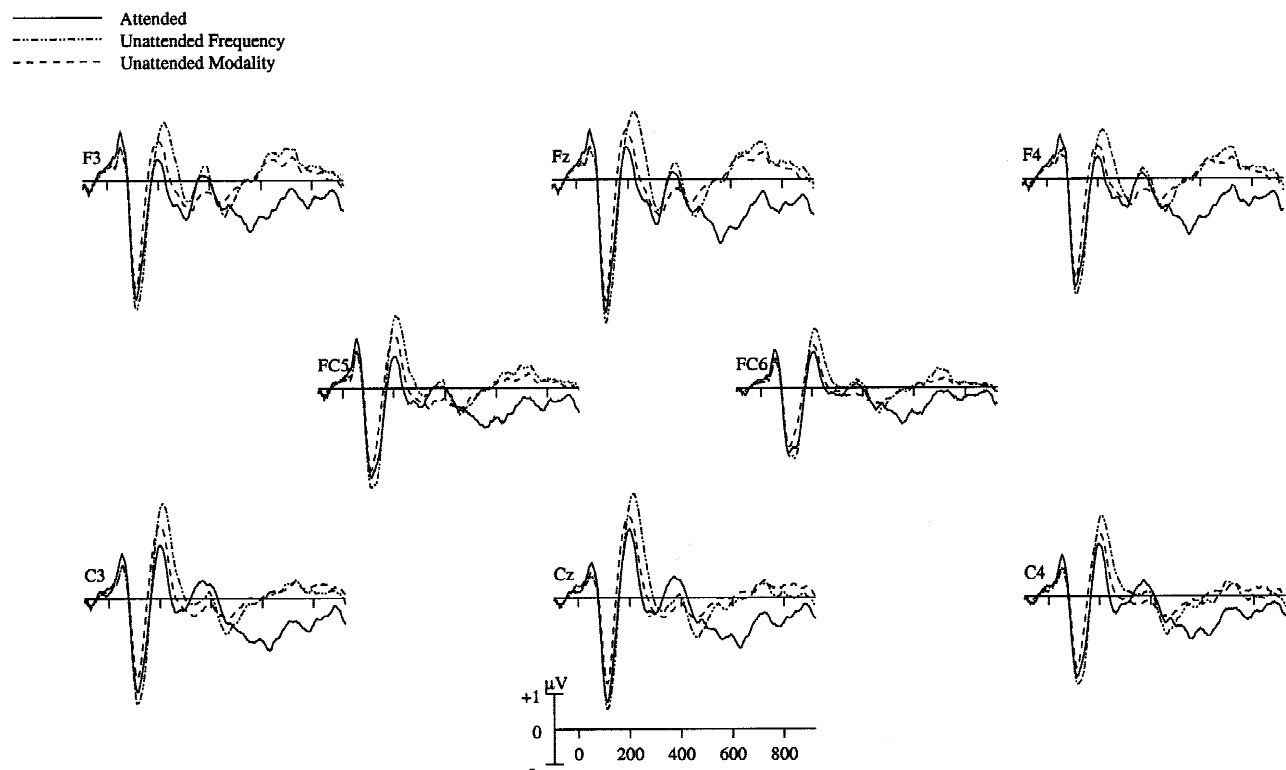


Figure 7. Grand average auditory ERP waveforms from a selection of eight anterior channels, obtained in the mixed modality condition. Intermodal attention is reflected in a reduced amplitude of the N1, which is followed by a pattern that is an intermediate of the attended frequency and unattended modality intramodal effects.

An additional finding of the present study (that also replicated the effects reported by De Ruiter et al. (1998)) was that intermodal attention had an earlier onset than the intramodal effect. ERP waveforms, in the attend visual conditions, differed significantly from those in the attend-auditory conditions over a period that started at about 170 ms, and lasted until over 300 ms after stimulus onset. This effect was obtained for ERPs to attended frequency as well as the unattended frequency stimuli. The relatively early onset of intermodal attention confirms our prior suggestion that stimuli that have a different modality are more distinct, and can thus be rejected in a relatively early phase of the selection process, as compared to stimuli that share the same modality.

For intermodal attention, the most notable effect in the ERP waveforms was a strong negative enhancement elicited by both attended frequency and unattended frequency stimuli at occipital sites, as compared to the ERP waveforms obtained in conditions when audition was relevant. This negative shift started directly after the peak of the N1 and lasted until about 300 ms after stimulus onset. A possible explanation for this intermodal effect is that it reflects an active inhibitory process that rejects visual stimuli, whenever audition is relevant (see De Ruiter et al., 1998, for a similar suggestion). If one assumes that the combined ERP waveforms elicited by attended frequency and unattended frequency stimuli can be taken as a baseline for visual processing, then the ERPs following unattended stimuli could be reminiscent of the late positivity that is elicited by unattended modality auditory stimuli and that was called "rejection positivity" by Alho,

Töttölä, Reinikainen, Sams, and Näätänen (1987). The idea that stimuli presented to unattended modalities are inhibited is also in line with the work of Skinner and Yingling (1977), who proposed that selective attention is implemented in the brain in terms of an active inhibition of everything that is irrelevant.

Localizing visual attention. For stimuli presented in the visual modality, intra- as well as intermodal attention became manifest in enhanced occipital selection negativities. Moreover, in accordance with prior studies, both types of attention showed almost identical topographical distributions (Alho et al., 1992; De Ruiter et al., 1998). The latter observations were further substantiated by dipole modeling of the ERP difference waves. As can be seen from Figures 3, 4, and 5, both intermodal and intermodal attention yielded similar solutions, with a symmetrical dipole pair being localized in the inferior occipital (extrastriate) areas. These areas coincide with the location of neural generators of occipital selection negativities that have been identified for visual attention to features such as spatial frequency (Heslenfeld et al., 1997) and color (Anllo-Vento, Luck & Hillyard, 1998; Lange et al., 1998).

Closer inspection of the dipole configurations shows that, although the visual difference waves all yielded dipoles with similar locations, their orientations differed considerably. A comparison of Figures 3 and 4 with Figure 5 illustrates this difference. One should note that the dipoles illustrated in Figures 3 and 4 both contain the ERP responses to attended frequency gratings in the subtraction, whereas those illustrated in Figure 5 do not. Hence, it can be argued

Table 3. Overview of the Time Slice Analyses for the Effects of Auditory Stimuli, as Performed on the Cz Electrode

Time (ms)	Mixed			Single	Single vs. mixed		
	af vs. uf	af vs. um	uf vs. um.	ATT	TASK	ATT	TASK \times ATT
60	NS	NS	NS	NS	9.63	NS	NS
72	NS	NS	NS	NS	12.99	NS	NS
84	NS	NS	NS	NS	NS	NS	NS
96	NS	NS	11.31	NS	NS	NS	NS
108	NS	NS	16.08	NS	NS	NS	NS
120	NS	NS	26.92	NS	NS	NS	NS
132	NS	9.02	35.12	NS	NS	NS	NS
144	NS	11.36	31.62	NS	NS	NS	NS
156	NS	9.79	19.20	NS	NS	NS	NS
168	NS	NS	NS	10.39	NS	NS	NS
180	NS	NS	NS	21.20	NS	11.92	NS
192	NS	NS	NS	31.88	NS	23.95	NS
204	15.99	NS	13.54	40.01	NS	41.17	12.22
216	26.13	NS	24.20	43.61	NS	55.22	10.91
228	32.01	NS	30.24	39.56	NS	53.67	NS
240	28.59	NS	31.26	30.88	NS	43.26	NS
252	21.58	NS	29.87	22.14	NS	32.69	NS
264	13.62	NS	22.87	15.31	NS	21.82	NS
276	NS	NS	15.03	10.05	NS	11.61	NS
288	NS	NS	NS	NS	NS	NS	NS
300	NS	NS	NS	NS	NS	NS	NS
312	NS	NS	NS	NS	NS	NS	NS
324	NS	NS	NS	NS	NS	NS	NS
336	NS	NS	NS	NS	NS	NS	NS
348	NS	NS	NS	NS	NS	NS	NS
360	NS	NS	NS	NS	NS	NS	NS
372	NS	NS	NS	NS	NS	NS	NS
384	NS	NS	NS	NS	NS	NS	NS
396	NS	NS	NS	NS	NS	NS	NS

Note: All reported F values (1,19 degrees of freedom) are significant with an $\alpha < .01$. NS: not significant; ATT: factor Attention; TASK: factor Task; af: attended frequency; uf: unattended frequency; um: unattended modality; Mixed: mixed modality condition; Single: single modality condition.

that the difference wave obtained by subtracting unattended modality gratings from attended frequency gratings contains both the intramodal selection negativity (from the attended frequency gratings) and the intermodal rejection positivity (from the unattended modality gratings, as discussed above). Following this line of thought, the difference wave obtained by subtracting unattended modality gratings from unattended frequency gratings then represents a purer estimate of the intermodal attention effect.

Data from the topographical analysis are also in line with this idea: unattended frequency and unattended modality gratings evoked similar scalp distributions, which both deviated from topography elicited by attended frequency gratings. Thus, it can be argued that subtracting unattended modality gratings from unattended frequency gratings yields the best estimate of intermodal attention. Therefore, a comparison of the dipole configurations for intramodal attention (as illustrated in Figure 3) with the dipole configuration for intermodal attention (as shown in Figure 5) shows similar locations, but markedly different orientations. The latter observation suggests that different neural circuits, both located within the extrastriate cortex, are responsible for intra- and intermodal attention.

As an alternative to the rejection positivity hypothesis postulated above, one could also argue that the unattended modality gratings (as operationally defined in the present study) represent the purest form of an unattended visual stimulus (i.e., not attended

on the basis of either modality or spatial frequency). If this were the case, differences between unattended modality gratings and unattended frequency gratings should not be interpreted as an active rejection process, suppressing visual stimuli when audition was relevant, but as a selection process, unique to selecting visual stimuli as such, when vision is the relevant modality. Attended frequency stimuli also show this negativity, which is shortly followed by an additional selection negativity, responsible for selecting visual stimuli with the attended spatial frequency. This interpretation also fits well within the hierarchical filter model of attention, as described by Heslenfeld (1998), and also explains why ERPs following attended frequency stimuli evoke a topographically different pattern of results than ERPs elicited by unattended frequency and unattended modality stimuli, because the ERPs generated by attended frequency stimuli contain one additional selection mechanism. According to LaBerge (1995), selective attention operates by enhancing the sensitivity of large groups of neurons, which are responsive to the attended stimulus feature. In terms of the present findings, this model could also explain the observed topographical differences. Intermodal attention would in this case represent the selective enhancement of the sensitivity of a large group of neurons, relatively early in the processing stream. The intramodal effect that follows could then reflect the additional enhanced sensitivity of neurons responsive to the attended spatial frequency of the attended frequency gratings.

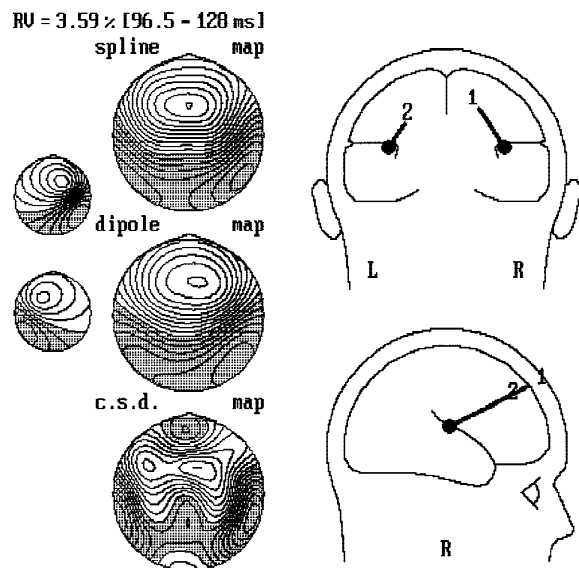


Figure 8. Dipole localization of the early (N1) auditory intermodal attention effect, as fitted on the difference wave obtained by subtracting ERPs following unattended modality stimuli from ERPs following attended modality stimuli. Contour maps were calculated at 107 ms after stimulus onset. Far left: scalp distribution resulting from dipoles 1 (top) and 2 (bottom). Left: measured scalp topography (top), modeled scalp topography (center), and current source density (CSD) map of the fitted dipoles. Contour spacing is $0.1 \mu\text{V}$ for the spline maps and $0.05 \mu\text{V}/\text{cm}^2$ for CSD maps. Right: coronal (top) and sagittal (bottom) view of the fitted dipole pair.

Auditory ERPs

It should be noted that the latency of the auditory ERP components (most notably the N1) appear to be somewhat delayed, compared to those latencies reported in previous studies (Hillyard et al., 1995). A possible difference is that in most studies auditory stimulation was done in a relatively direct fashion, that is, using a headphone set. In the present study, a more indirect form of stimulation was used. Combined with a relatively fast presentation rate and the presence of visual stimuli, this could have led to a somewhat slower orientation of the auditory stimuli. Visual inspection of the data obtained by Eimer & Schröger (1998), who also used this method, also suggests that the auditory N1 latencies are somewhat delayed in their study.

Intramodal attention. Intramodal auditory attention became manifest in a negative modulation of ERP waveforms in the area between about 200 and 300 ms. This late process bears resemblance to the later part of the processing negativity that is assumed to reflect maintenance and rehearsal of the "attentional trace" or some form of further processing of the attended stimuli (Näätänen, 1982, 1990). The late emergence of processing negativity in the current study could be due to the fact that stimuli were presented at a relatively low rate. Experiments in which attended and unattended stimuli were delivered at very fast rates have usually reported much earlier attentional-related modulations (Hillyard et al., 1995; Woldorff & Hillyard, 1991). Another factor that might have contributed to the late emergence of the processing negativity is that in the present study, intramodal attention required a highly precise stimulus selection, namely on the basis of modality, pitch, and duration of stimuli.

The later part of the processing negativity is usually manifested as a frontally distributed component, whereas the intramodal at-

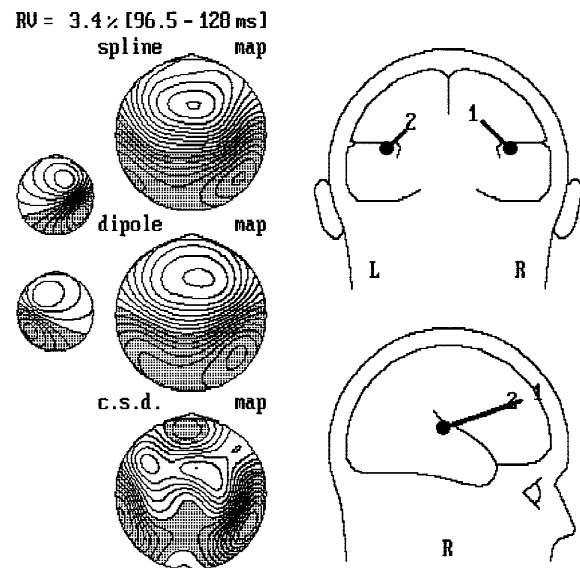


Figure 9. Dipole localization of the early (N1) intermodal auditory attention effect, as fitted on the difference wave obtained by subtracting unattended modality from unattended frequency ERPs. Contour maps were calculated at 107 ms after stimulus onset. Far left: scalp distribution resulting from dipoles 1 (top) and 2 (bottom). Immediate left: measured scalp topography (top), modeled scalp topography (center), and current source density (CSD) map of the fitted dipoles. Contour spacing is $0.1 \mu\text{V}$ for the spline maps and $0.05 \mu\text{V}/\text{cm}^2$ for CSD maps. Right: coronal (top) and sagittal (bottom) view of the fitted dipole pair.

tention effect in the present study was distributed more broadly across the midline (see also Figures 6 and 7). Note that processing negativities that are elicited in intermodal auditory task usually have a more posterior focus (Alho et al., 1992, 1994; Woods et al., 1992). Because the subjects received mixed streams of visual and auditory stimuli, it is possible that intermodal control processes (and related generators) also contributed to some extent to the late negative deflections that were elicited in the intramodal condition.

Intermodal attention. Prior intermodal studies have established that intermodal auditory attention is manifested in a biphasic modulation of early components (enhancement of N1 as well as of P2 amplitude). This pattern contrasts with the typical monophasic processing negativities found in intramodal auditory conditions (Alho et al., 1992; Woods et al., 1992).

Our results partially replicate these earlier findings. Tones presented when audition was relevant elicited larger N1 components than tones presented in the attend visual condition. However, for the subsequent P2 region, a more complex pattern of attentional modulation was found.

In this region, effects of intermodal attention on the ERP waveforms strongly depended on the type of auditory stimuli that were presented when audition was the relevant modality. For unattended modality tones, intermodal attention was manifested in a positive displacement that affected primarily the descending limb of P2. This effect clearly resembles the typical P2 enhancement (also labeled as Pd, i.e. positive difference) effect reported in earlier intermodal studies (e.g., Alho et al., 1992; Woods et al., 1992).

However, for the attended frequency tones, a different pattern was found. These stimuli elicited a greater negativity (hence smaller P2s) than tones presented in conditions where vision was relevant.

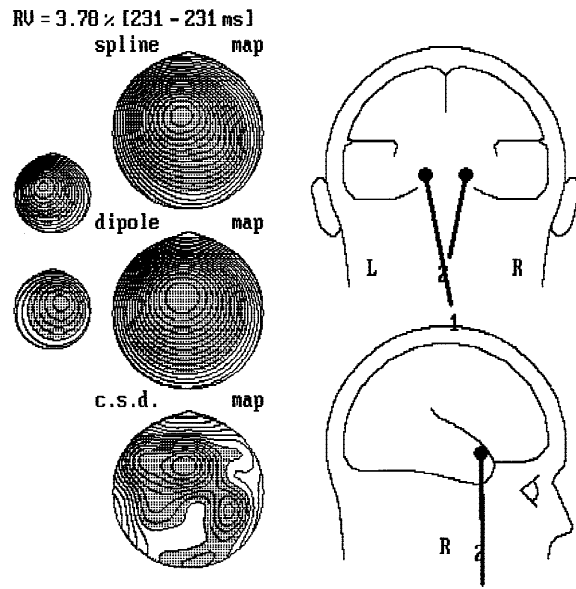


Figure 10. Dipole localization of the late (P2) effect of auditory intra-modal attention, as fitted on the difference wave obtained by subtracting ERPs following unattended frequency stimuli from ERPs following attended frequency stimuli. Contour maps were calculated at 321 ms after stimulus onset. Far left: scalp distribution resulting from dipoles 1 (top) and 2 (bottom). Immediate left: measured scalp topography (top), modeled scalp topography (center), and current source density (CSD) map of the fitted dipoles. Contour spacing is $0.1 \mu\text{V}$ for the spline maps and $0.05 \mu\text{V}/\text{cm}^2$ for CSD maps. Right: coronal (top) and sagittal (bottom) view of the fitted dipole pair.

It is possible that the latter effect resulted from overlap between P2 and processing negativity (in particular the later component, Ndl). As noted earlier, in the current study, attended frequency stimuli required a precise selection that could have elicited a strong processing negativity. Following the same reasoning, it can be argued that unattended frequency tones elicited smaller processing negativities because these stimuli need a less precise selection.

We also like to draw attention to a similar effect of intermodal auditory attention reported by Alho et al. (1994, Figure 1). In this experiment, subjects also had to discriminate between two levels of auditory attention within the attended modality (i.e., attended and unattended ears). In contrast with their prior studies (in which they reported a positive displacement of ERPs to attend auditory stimuli) they now showed that ERPs to attended frequency stimuli were negatively displaced relative to ERPs elicited by auditory stimuli when the visual modality was relevant.

The present results clearly demonstrate that intermodal attention is expressed differently in ERP waveforms to visual and auditory stimuli. This is most conspicuous for stimuli that had to be rejected on the basis of their modality. When audition was the relevant modality, visual stimuli did not elicit any selection negativities. In contrast, when vision was the relevant modality auditory stimuli elicited ERPs that were somewhat intermediate between ERPs to the attended frequency and unattended frequency auditory stimuli (see also Michie, Solowij, Crawford, & Clue, 1993, for similar results of auditory intermodal attention).

It is interesting to speculate about the possible reasons of these deviating ERP patterns. One possibility is that the ERP findings reflected fundamental differences in mechanisms underlying se-

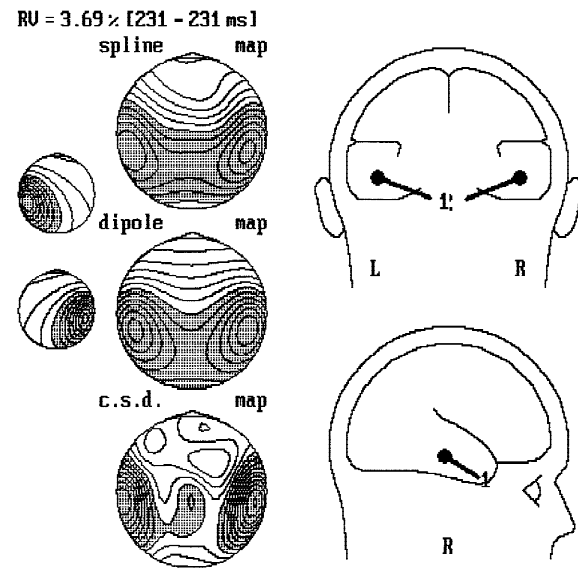


Figure 11. Dipole localization of the late (P2) effect of auditory inter-modal attention, as fitted on the difference wave obtained by subtracting unattended modality from unattended frequency ERPs. Contour maps were calculated at 321 ms after stimulus onset. Far left: scalp distribution resulting from dipoles 1 (top) and 2 (bottom). Immediate left: measured scalp topography (top), modeled scalp topography (center), and current source density (CSD) map of the fitted dipoles. Contour spacing is $0.1 \mu\text{V}$ for the spline maps and $0.05 \mu\text{V}/\text{cm}^2$ for CSD maps. Right: coronal (top) and sagittal (bottom) view of the fitted dipole pair.

lection of visual and auditory information. That is, auditory stimuli could be more difficult to be rejected or “filtered out” on the basis of their modality than visual stimuli (Schröger, 1996).

Several factors could have contributed to these effects. First, auditory stimuli have the reputation of eliciting strong representations in sensory memory (e.g., Näätänen, 1990). Second, in the present experiment, stimuli were presented at a relatively low delivery rate. Thus, frequently presented auditory stimuli, which the subject was instructed to ignore, may still have elicited some residual processing negativity. A third factor could have been that the visual task was somewhat easier to perform than the auditory task (see also Table 1). This could have resulted in some “spare” capacity left to be captured by auditory stimuli when subjects paid attention to stimuli in the visual modality.

Another way of putting it is that when subjects attended to visual stimuli, auditory stimuli still covertly attracted attention. Note also that the analysis of behavioral results (Table 1) did not give rise to suspicion that subjects suffered more from interference by auditory stimuli in the visual task than vice versa. This fits well with the idea that covert attention shifts do not necessarily have to interfere with controlled attentional processes.

Localizing auditory attention. A satisfactory dipole fit was obtained for the early effect of intermodal attention, that is, the negative modulation of the N1 (between 110 and 130 ms) component. As can be seen in Figures 8 and 9, dipole pairs fitted to both attended frequency minus unattended frequency as well as unattended frequency minus unattended modality difference waves yielded highly similar solutions, which suggests the presence of neural sources in the supratemporal planes of the auditory cortex. The orientations of the dipoles are also in accordance with the

frontocentral distribution of the intermodal attention effect in the ERPs in the N1 area.

The location of the neural generators underlying intermodal attention seem to correspond closely to the location of generators that have been identified for intramodal auditory attention. These generators have been reported in dichotic listening tasks for the early Nd component that overlaps with the N1 from neuromagnetic studies (Woldorff, Hackley, & Hillyard, 1993). These authors have further demonstrated that their "M100" effect inverted in polarity at the anterior relative to the posterior scalp sites, which is consistent with a vertically oriented current dipole located in the supratemporal plane.

The later stages of auditory attention yielded only reasonable dipole fits for one intramodal (attended frequency minus unattended frequency) and one intermodal (unattended frequency minus unattended modality) difference wave. As can be seen from Figures 10 and 11, these solutions yielded markedly different results. Effects of intramodal attention could be localized in the inferior medial frontal areas, as seen from Figure 10. Note that this solution is similar to results reported by Benedict et al. (1998), who found increases in regional cerebral blood flow in both left and right anterior cingulate as well as in the right prefrontal cortex, during an auditory attention task. Possibly this dipole reflects the later stages of a consciously controlled, higher order target detection system (e.g., LaBerge, 1990; Posner & Petersen, 1990).

The late intermodal attention effect, on the other hand, (Figure 11) yielded a substantially different dipole configuration. These locations bear some resemblance to the dipole configurations reported for the earlier effects and could indicate a possible involvement of the auditory cortex during the later stages of auditory selection. Note, however, that the orientations of the later dipoles are markedly different from those reported for the earlier latencies.

Single Versus Mixed Modality Conditions

The analysis of behavioral measures (errors and response times) suggests that the auditory task was somewhat more difficult than the visual tasks. No evidence was found, however, that performance in the mixed conditions was worse than in the single modality conditions. This suggests that when subjects focused their attention on stimuli presented in the relevant modality, they were not affected in their performance by interference by stimuli belonging to the other modality.

However, the analysis of ERP results yielded a somewhat different and more complex pattern of results that depended on the modality of the stimuli. Interestingly, the mode of presentation of stimuli (single versus mixed) affected the same ERP components that were shown to be sensitive to (intra- and intermodal) attentional manipulations. In the visual modality, stimuli in the mixed condition elicited enhanced P2 components in the early region between 100 to 200 ms relative to stimuli in the single modality condition. Likewise, auditory stimuli in the mixed modality condition elicited larger N1 (Nde) components and larger processing negativity (Ndl) in the region between 200 to 300 ms than auditory stimuli in the single condition. Note that although the observed amplitude differences between the single and mixed modality conditions did not influence the effects of intramodal attention on visual stimuli, such an interaction between task and attention was present for the auditory stimuli. For the latter, the amplitude difference between attended frequency and unattended modality stimuli was somewhat enlarged in the single modality condition as compared to the mixed modality condition.

We propose the following interpretation of these findings. It is possible that in the mixed condition, the intermodal attention effect partly overlapped with the intramodal effect. This holds in particular for ERPs to attended frequency stimuli, which could have absorbed effects of intra- as well as intermodal attention. Because both forms of attention globally affect the same ERP components (i.e., selection and processing negativities, for the visual and auditory attention conditions), it seems plausible that their joint effects caused a further enhancement of the attention-related negativities.

An alternative interpretation of these findings (that is not necessarily in conflict with the previous interpretation) is that the single versus mixed modality effect could have resulted from the higher presentation rate of stimuli. Recall that in the mixed modality condition, subjects received on the average about twice as many stimuli per time unit as in the single modality condition. Thus, although the distributions and ISIs of the visual and auditory stimuli were approximately the same in the single and mixed conditions, it is still possible that the higher temporal density of stimuli in the bimodal stimulus sequences caused a higher stimulus load and hence a stronger "stimulus set". This in turn could have caused an enhancement of early activity in selected sites of extrastriate (visual modality) or auditory (auditory modality) cortex. The latter hypothesis is also consistent with recent attentional models (LaBerge, 1995; Lavie, 1995; Lavie & Tsal, 1994; Näätänen, 1990). Another finding that is worth mentioning was that for the visual modality, the early effect of intramodal visual attention was seen much less clearly in the mixed than in the single modality condition. This can perhaps also be explained in terms of a "ceiling effect" in the mixed visual condition that was due to the summation of intra- and intermodal attention.

General Conclusions

The present data are well in agreement with an hierarchical model of selective attention (Hansen & Hillyard, 1983). The relatively early onset of intermodal attention, which was found for both visual and auditory stimuli, indicates that modality is most likely selected prior to stimulus-relevant features within the attended modality. In terms of a neurophysiological model of attention (LaBerge, 1990, 1995; Posner & Petersen, 1990), the early intermodal effects probably reflect a general sensory enhancement, which occurs relatively early in the processing stream. It should be noted that the present data indicate the presence of such an enhancement mechanism for both visual and auditory data.

The intramodal effect, on the other hand, differed between visual and auditory stimuli. Whereas for visual stimuli evidence was found for the possible involvement of the visual cortex in visual intramodal attention, for auditory attention this was not the case. A possible reason why the present study failed in identifying reasonably plausible neural sources for intramodal attention could be that the brain activity generated by auditory stimuli is much more diffuse than that generated by visual stimuli. This could also be evidenced by the observation that the residual variance of dipole models that included latencies closely surrounding the N1 peak was markedly increased, indicating the existence of a more complex pattern of neural generators active around the same time as the N1 generators in the auditory cortex. Given the relatively small number of electrodes available in the present study, further research is required to unravel the precise composition of these dipole patterns.

The neural sources obtained for intramodal visual attention were localized in a somewhat anterior part of the visual areas. In terms of a neurophysiological explanation, it is possible that these dipoles represent the enhancement of groups of neurons specifically sensitive to spatial frequencies. This interpretation is in line with PET findings from Corbetta, Miezin, Dobmeyer, Shulman, and Petersen (1991) and Heinze et al. (1994), who have demonstrated the existence of bilateral areas in the fusiform gyrus and parahippocampal gyri, which were selectively responsive to shape. It is thus possible that visual attention

consists of a general enhancement of responsiveness of the visual areas, combined with a specific enhancement in sensitivity of neurons responsive to the attended stimulus feature. If this is the case, then the difference in the onset of the inter- versus intramodal visual attention effect possibly reflects the time needed to transmit visual stimuli from the general (primary) visual areas to the secondary visual areas responsive to the attended stimulus feature. It is not until visual stimuli reach the feature-specific areas in the extrastriate cortex that the neurons there can respond to them.

REFERENCES

- Alho, K., Töttölä, L., Reinikainen, K., Sams, M., & Näätänen, R. (1987). Brain mechanism of selective listening reflected by event-related potentials. *Electroencephalography and Clinical Neurophysiology*, 68, 458–470.
- Alho, K., Woods, D. L., & Algazi, A. (1994). Processing of auditory stimuli during auditory and visual attention as revealed by event-related potentials. *Psychophysiology*, 31, 469–479.
- Alho, K., Woods, D. L., Algazi, A., & Näätänen, R. (1992). Intermodal selective attention II. effects of attentional load on processing auditory and visual stimuli in central space. *Electroencephalography and Clinical Neurophysiology*, 82, 356–368.
- Anllo-Vento, L., Luck, S. J., & Hillyard, S. A. (1998). Spatio-temporal dynamics of attention to color: Evidence from human electrophysiology. *Human Brain Mapping*, 6, 216–238.
- Bastiaansen, M. C. M., Brunia, C. H. M., & Böcker, K. B. E. (1999). ERD as an index of anticipatory behavior. In G. Pfurtscheller & F. Lopez da Silva (Eds.), *Event-related desynchronization: Handbook of electroencephalography and clinical neurophysiology*, revised series (Vol. 6, pp. 203–217). Amsterdam: Elsevier.
- Benedict, R. H. B., Lockwood, A. H., Schucard, J. L., Schucard, D. W., Wack, D., & Murphy, B. W. (1998). Functional neuroimaging of attention in the auditory modality. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience*, 9(1), 121–126.
- Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1991). Selective and divided attention during visual discriminations of shape, color, and speed: Functional anatomy by positron emission tomography. *Journal of Neuroscience*, 11, 2383–2402.
- De Ruiter, M. B., Kok, A., & van der Schoot, M. (1998). Effects of inter- and intramodal selective attention to non-spatial visual stimuli: An event-related potential analysis. *Biological Psychology*, 49, 269–294.
- Eimer, M., & Schröger, E. (1998). Effects of intermodal attention and cross-modal attention in spatial attention. *Psychophysiology*, 35, 313–327.
- Hackley, S. A., Woldorff, M., & Hillyard, S. A. (1990). Cross-modal selective attention effects on retinal, myogenic, brainstem, and cerebral evoked potentials. *Psychophysiology*, 27, 195–208.
- Hansen, J. C., & Hillyard, S. A. (1983). Selective attention to multidimensional auditory stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 1–19.
- Heinze, H., Mangun, G. R., Burchert, W., Hinrichs, H., Scholz, M., Münte, T. F., Göss, A., Scherg, M., Johannes, S., Hundeshagen, H., Gazzaniga, M. S., & Hillyard, S. A. (1994). Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature*, 372, 543–546.
- Heslenfeld, D. (1998). *Features and attention in vision: An analysis of electromagnetic brain responses*. Unpublished doctoral dissertation, University of Amsterdam.
- Heslenfeld, D. J., Kenemans, J. L., Kok, A., & Molenaar, P. C. M. (1997). Feature processing and attention in the human visual system: An overview. *Biological Psychology*, 45, 183–215.
- Hillyard, S. A., Mangun, G. R., Woldorff, M. G., & Luck, S. J. (1995). Neural systems mediating selective attention. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 661–681). Cambridge MA: MIT Press.
- Huizenga, H. M., & Molenaar, P. C. M. (1994). Estimation and testing the sources of evoked potentials in the brain. *Multivariate Behavioral Research*, 29, 237–262.
- Kenemans, J. L., Baas, J. M. P., Mangun, G. R., Lijffijt, M., & Verbaten, M. N. (2000). On the processing of spatial frequencies as revealed by evoked-potential source modeling. *Clinical Neurophysiology*, 111, 1113–1123.
- Kenemans, J. L., Kok, A., & Smulders, F. T. Y. (1993). Event related potentials to conjunctions of spatial frequency and orientation as a function of stimulus parameters and response requirements. *Electroencephalography and Clinical Neurophysiology*, 88, 51–63.
- Kenemans, J. L., Molenaar, P. C. M., & Verbaten, M. N. (1991). Models for estimation and removal of artifacts in biological signals. In R. Weiskunat (Ed.), *Digital biosignal processing*. New York, NY: Elsevier.
- LaBerge, D. (1990). Thalamic and cortical mechanisms of attention suggested by recent positron emission tomographic experiments. *Journal of Cognitive Neuroscience*, 2, 358–372.
- LaBerge, D. (1995). *Attentional processing: The brain's art of mindfulness*. Cambridge, MA: Harvard University Press.
- Lange, J. J., Wijers, A. A., Mulder, L. J. M., & Mulder, G. (1998). Color selection and location selection in ERPs: Differences, similarities and 'neural specificity'. *Biological Psychology*, 48, 153–182.
- Lange, J. J., Wijers, A. A., Mulder, L. J. M., & Mulder, G. (1999). ERP effects of spatial attention and display search with unilateral and bilateral stimulus displays. *Biological Psychology*, 50, 203–233.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 451–468.
- Lavie, N., & Tsai, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception and Psychophysics*, 56, 183–197.
- McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, 62, 203–208.
- Michie, P. T., Solowij, N., Crawford, J. M., & Clue, L. C. (1993). The effects of between source discriminability on attended and unattended auditory ERPs. *Psychophysiology*, 30, 205–220.
- Mountcastle, V. B. (1978). Brain mechanisms of directed attention. *Journal of the Royal Society of Medicine*, 71, 14–27.
- Näätänen, R. (1982). Processing negativity: An evoked potential reflection of selective attention. *Psychological Bulletin*, 92, 605–640.
- Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, 13, 201–288.
- Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Petersen, S. E., Robinson, D. L., & Morris, J. D. (1987). Contributions of the thalamus to visual spatial attention. *Neuropsychology*, 25, 97–107.
- Pick, H. L., Warren, D. H., & Hay, J. C. (1969). Sensory conflict in judgments of spatial direction. *Perception and Psychophysiology*, 6, 203–205.
- Platt, B. B., & Warren, D. H. (1972). Auditory localization: The importance of eye movements and a textured visual environment. *Perception and Psychophysiology*, 12, 245–248.
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, 13, 25–42.
- Previc, F. H., & Harter, M. R. (1982). Electrophysiological and behavioral indicators of selective attention to multifeature gratings. *Perception and Psychophysics*, 32, 465–472.
- Scherg, M., & Berg, P. (1995). *BESA brain electric source analysis. User manual*. Version 2.1. Munich.
- Schröger, E. (1996). A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *Journal of Cognitive Neuroscience*, 8, 528–539.

- Skinner, J. E., & Yingling, C. D. (1977). Central gating mechanisms that regulate event-related potentials and behavior. In J. E. Desmedt (Ed.), *Attention, voluntary contraction and event-related cerebral potentials: Progress in clinical neurophysiology* (Vol. 1, pp. 30–69). Basel: Karger.
- Spence, C. J., & Driver, J. (1997). On measuring selective attention to an expected sensory modality. *Perception and Psychophysics*, 59, 389–403.
- Van der Stelt, O., Kok, A., Smulders, F. T. Y., Snel, J., & Gunning, B. (1998). Cerebral event-related potentials associated with selective attention to color: Developmental changes from childhood to adulthood. *Psychophysiology*, 35, 227–239.
- Wijers, A. A., Lange, J. J., Mulder, G., & Mulder, L. J. M. (1997). An ERP study of visual spatial attention and letter target detection for isoluminant and nonisoluminant stimuli. *Psychophysiology*, 34, 553–565.
- Wijers, A. A., van Hooff, H., Lange, J. J., Peters, M., & Dunajski, Z. (1993). Topography and source analysis of brain activity associated with selective spatial attention and memory search. *Brain Topography*, 5, 383–388.
- Woldorff, M. G. (1993). Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction. *Psychophysiology*, 30, 98–119.
- Woldorff, M. G., Hackley, S. A., & Hillyard, S. A. (1993). The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones. *Psychophysiology*, 28, 30–42.
- Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology*, 79, 170–191.
- Woods, D. L., Alho, K., & Algazi, A. (1992). Intermodal selective attention. I: Effects on event-related potentials to lateralized auditory and visual stimuli. *Electroencephalography and Clinical Neurophysiology*, 82, 341–355.
- Wurtz, R. H., Goldberg, M. E., & Robinson, D. L. (1980). Behavioral modulation of visual responses in the monkey. In J. M. Sprague & A. Epstein (Eds.), *Progress in psychobiology and physiological psychology* (Vol. 9, pp. 43–83). New York, NY: Academic Press.

(RECEIVED June 21, 2000; ACCEPTED January 6, 2001)